区块链技术在 UGC 内容审核方面的研究与实践

张福成 郭志帅 肖 勇
（东软集团，辽宁 沈阳 110179）

摘 要：区块链技术作为我国核心技术自主创新的重要突破口，在媒体深度融合领域必将起到举足轻重的作用。本文分析了 UGC 内容审核风险控制的必要性以及区块链技术体系的特点和 UGC 内容风险控制目前面临的问题，结合基于区块链技术的内容审核标准阐述了基于区块链技术的 UGC 内容审核风险控制系统的建设思路和建设重点。并基于系统的建设实践明确了系统功能架构设计，旨在通过区块链技术的实践和探索来保障新闻内容的安全生产。

关键词：内容审核；区块链；数据可信；全流程追溯；平台化 **中图分类号：**G712 **文献标识码：**A
文章编号：1671-0134（2021）11-020-04 DOI：10.19483/j.cnki.11-4653/n.2021.11.005
本文著录格式：张福成，郭志帅，肖勇 . 区块链技术在 UGC 内容审核方面的研究与实践 [J]. 中国传媒科技，2021（11）：20-23.

2018 年，随着区块链技术的发展，在国内外媒体行业内也掀起了一场区块链应用热潮，业界纷纷探索在媒体融合领域中区块链技术能有哪些应用价值，尝试通过区块链技术来助力媒体融合发展的深度和广度。区块链作为分布式数据库系统，具有多节点参与数据存储的特点，区块链中的任意一个节点通过一致性算法来存储相同的数据，保证了存储的安全可靠。其具有的去中心化、难以篡改、扩展性大、灵活性强等技术特征在数字内容版权保护、信息加密等领域给出了全新的解决方案和重要支撑。区块链是一种包含了哈希算法、非对称加密技术、数字签名技术、时间戳技术、数字证书的复合技术，通过区块链技术存储数字摘要，使数字内容不可伪造以及篡改，达到存证目的，从而为鉴权、确认、追溯等业务提供可信的数据支撑。区块链技术在打击虚假新闻、内容审核、版权保护、广告精准投放等方面为传媒行业带来新的可能。[1] 凭借颠覆式创新的计算模式，区块链技术在新一代信息技术基础上形成了新一代融媒体技术，引领国内传统媒体和新媒体融合的未来走向。[2]

UGC，全称为 User Generated Content，是用户生产内容、用户原创内容的简称。UGC 起步于互联网领域，即互联网用户将自己制作、自己原创的内容在互联网平台上展示以及提供给其他的互联网用户。UGC 并非某种具体的业务，其本质是用户使用互联网的一种交互方式，由以前单一的信息获取、数据下载变成了内容提供、内容上传。随着我国移动互联网运用的发展，移动互联网网络用户的交互作用得到了进一步体现，用户既是内容的浏览者，更是移动互联网内容的创造者，尤其是以微信订阅号、抖音号、头条号、主流新闻客户端入驻号等各类官媒以及自媒体平台为代表，UGC 内容已经在内容传播、内容服务、电子商务等方面产生了重要的影响。

1.UGC 内容风控面临的挑战

我国相关部门和决策者针对网络内容合规的管控一直都执行着一套非常严格的标准，尤其 5G 技术的出现更是使内容呈现出海量性、即时性等特征，使得内容监管工作颇具挑战。[3] 因此近年也出现监管标准越来越细、惩罚力度越来越大的趋势。UGC 作为以用户自主生产内容为核心业务的产品，自然是内容风险重灾区。对存在违规现象的平台或者账户，轻则点名约谈警告，重则无限期关停。对平台管理方，尤其是传统的报社、电视台等国家新闻传媒机构，在进军互联网占领主战场的过程中，也在不断建设和尝试 UGC 内容平台运营，而内容风险控制是关乎到平台生死存亡的大事，在用户体验、平台运营策略等多方面都要基于内容审核风险控制为基础来进行取舍。UGC 内容的几个主要风险因素：

● 涉及违法违规的：违法违规是所有风险项目中最为严重的，是审核风控的底线，需要用最严格的手段控制，否则将影响到平台生存。

● 影响用户体验和平台氛围的：譬如辱骂、带有攻击色彩的信息、无意义的灌水文字，以及广告等内容。此类内容风险适中。

● 影响平台内容方向的：平台通常都有自己的内容侧重，尤其是报业、广电等传媒机构，在政治性、严肃性等方面都有自己的相关要求和准则。无关正确舆论方向的内容不应该得到鼓励，否则会影响品牌形象，例如过多存在一些八卦娱乐信息，势必会对专业度有负面冲击，但此类内容风险较低。

● 涉及用户安全的：诸如诈骗、传销等灰色产业信息，作为内容平台有义务保障用户财产等各方面安全，同时也是平台严谨性、权威性的侧面体现，这类内容也需要重点关注。

● 版权侵犯：信息技术的发展，使得数字版权侵权

现象加剧，版权信息关乎到了平台的内容合法运营，使用或者传播了侵权内容，存在针对平台的索赔风险，因此此类也是需要重点审核的内容。

2. 区块链技术的核心特性

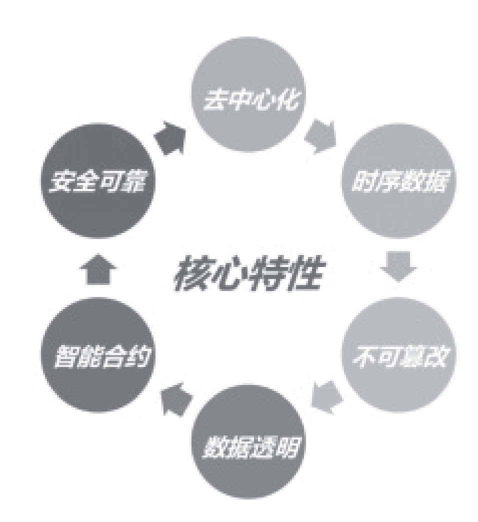


图 1 区块链技术核心特性

●安全可靠特性：单一节点故障不影响系统服务，密码学保护数据安全。

● 去中心化特性：没有中心服务器，数据使用分布式方式存储，系统具备高冗余性。

● 时序数据特性：数据被加盖时间戳，按照时间顺序构成区块链。

● 不可篡改特性：修改数据必须修改多数节点，数据结构与数据内容直接相关，修改历史数据会影响链式结构。

● 数据透明特性：数据以及数据执行规则对所有相关方透明。

● 智能合约特性：程序执行业务条款、业务在各节点同时执行、合约内容对相关方公开。

3. 基于区块链技术的内容审核发展状况

2021 年 4 月 2 日，国家广电总局正式发布了《基于区块链的内容审核标准体系（2021 版）》，该标准体系的核心目标是基于区块链的内容审核标准体系促进和提升媒体内容的健康、可持续发展。标准针对以区块链为核心的内容审核的全流程进行了规范化定义，涵盖了业务处理流程、管理过程、安全策略等各个环节，旨在通过标准来指导技术的应用以及平台的建设。

整个标准体系从系统、业务、安全、管理四个方面进行了详细的标准细则阐述。

● 系统类标准。针对区块链内容审核的系统架构、数据格式、系统接口等技术要求进行了规范化的说明，包括总体系统技术规范、媒体内容的分类以及标志、媒体内容存储模式以及共享模式、区块单元的数据详细格式、系统之间的接口、跨链交互的技术规范等标准。

● 业务类标准。从区块链内容审核的业务流程出发，定义了涵盖内容审核、溯源以及重审等环节的各项标准。

● 安全类标准。从审核安全角度出发，以系统安全技术要求和安全管理审计规范为核心，规范了基于区块链内容审核安全的各项标准。

● 管理类标准。从监管角度出发，详细规范了监管、内容审核评价、系统、节点评估等标准，旨在将基于区块链内容审核的监管要求进行规范化定义。

随着标准的出台，传媒单位、技术公司逐步开始了基于区块链技术的内容审核平台建设的尝试和探索，从而为标准的落地、有效实施提供实践依据和宝贵经验。通过基于区块链技术的审核平台建设，可以对审核过程、审核结果进行全流程的记录，为后续的侵权纠纷、内容追责提供具备公信力的证据，不论是对内容生产者还是内容审核方都具备更强的约束力，从而保障和提升内容审核严谨性，从而推动内容生产安全性的不断提升。

4. 基于区块链技术的 UGC 内容审核平台建设要点分析

4.1 技术架构选定

区块链是一项技术，但同时也是一种架构应用，因此区块链技术应用过程中架构的实现是技术体系的核心。区块链技术体系包括网络层、共识层、数据层、智能合约层和应用层，每一层具体的含义本文中不做过多的说明。UGC 内容审核风控系统的业务功能主要是基于应用层开展建设。应用层主要包括了与区块链底层交互的智能合约以及对应的 UGC 内容风控前端 WEB 应用。WEB 应用系统通过接口调用智能合约，根据业务规则进行数据存取，从而完成相应的业务功能。

区块链的基础建设可有两种方式，一个是公有云服务、一个是自有链搭建。公有云服务适合开发技术能力一般的组织机构选择使用，直接利用公有云的区块链进行业务应用的建设，打造自己的区块链业务支撑，但需要按照时长或者用量支付服务费用。对于具备较好技术研发、维护能力的组织机构可以自行搭建区块链基础平台，从而平台的可控性、自主性方面更具优势。对于基于区块链的 UGC 内容审核风控系统的底层区块链，需要结合建设方自身特点进行灵活选择。

4.2 审核数据链上加密存储

UGC 内容基于自身特点具备海量、零散、种类众多的特点，针对审核过程需要将 UGC 内容本身信息、审核流程信息、审核结果信息进行上链存储，从而保障 UGC 内容、审核过程、审核结果的数据完整性，为后续的审核全过程追溯提供数据支撑。

● UGC 内容本身信息：系统内唯一标识、发布人、发布时间、发布渠道、发布栏目、发布节点等。

●审核过程信息：审核的种类、审核起止时间、审核操作人、过程时长等。

●审核的结果信息：审核结果、审核风险标识、审

核等级、预警分类等。

审核关联信息众多，全部上链会导致存储空间占用过多以及计算过程耗时的问题，因此可以将各部分信息进行哈希计算，仅仅将哈希码上链实现加密存储即可。

4.3 去中心化的共识机制保证了防篡改的全流程追溯

“区块链的去中心化以及分布式数据存储可以帮助新闻的发表认证更加严密、每个节点既独立存储，又不可被人为地抹除。”[4] 同理，基于区块链去中心化的共识机制保证了内容审核全流程的各个环节数据都无法进行恶意篡改。针对任何一条内容信息，业务系统通过接口封装，基于唯一识别编码从链上取得相应信息，对信息进行解析，从而展示出内容信息本身以及关联的审核信息。如果发生恶意篡改，将会导致链上存储的哈希码同被改后的数据生成的哈希码存在差异。系统实现层面，审核对象以及审核信息均在业务系统内存储，对应的哈希码将上链，因此对业务系统内数据进行的篡改将导致哈希值与链上的不一致，从而发现篡改事实。审核平台重在记录审核的可信全流程，通过一审、二审等流程对UGC内容进行是否存在违规的判定，并将每次审核的数据存储到区块链之上，达成不可篡改的目的。

4.4 多类型内容审核业务流程确立

传统的内容审核业务流程比较单一，针对数据进行审核，并形成审核结果，记录到业务系统的数据库之中，数据库常以关系型数据库为主。内容审核结合了区块链技术后，需要针对不同类型数据的审核制定比较完善的审核业务流程，从而才能明确区块链技术在业务各个环节的定位以及运用方式和应该发挥的作用。

业务流程规划示例：

职业生产内容（OGC）：内容生产 –> 内容审核（三审三校）–> 审核内容、审核过程、审核节点的哈希码上链 –> 成品后审 –> 后审结果的哈希码上链。

用户生产内容（UGC）：内容生产 –> 机器审核 –> 审核内容、审核结果的哈希码上链 –> 人工复审 –> 审核人、审核结果的哈希码上链。

另外针对图片、视频、文本等多种类型业务数据在存储上也要基于实际业务形态来划分不同的流程策略，譬如视频存取关键帧、文本存储关键词等策略，从而控制上链信息数量以及提升查询等应用的性能。

4.5 审核信息的跨链共享机制的确立

在区块链所面临的诸多问题中，区块链之间互通性极大程度的限制了区块链的应用空间。不论对于公有链还是私有链，跨链技术是实现区块链互联互通、提升区块链互操作性与可扩展性的重要技术手段。[5]

区块链的跨链应用需要解决“如何确认”“在哪确认”“由谁来确认”几个关键问题，通常采用公证人机制（Notary Schemes）、侧链/中继（Sidechains/Relays）、哈希锁定（Hash-locking）、分布式私钥控制（Distributed Private Key Control）、原子交换（Atomic Swap）等几种实现方式。

内容审核业务平台跨链信息共享的核心在于由谁来进行信息在不同链之间的互信认证以及被审核数据和结果数据的共享，因此针对内容审核业务的跨链共享可通过建立跨链服务中心来解决。服务中心提供如下核心功能：

- 权限认证：所有机构均需要在服务中心进行注册并分配认证秘钥，所有系统间数据共享接口均需要通过秘钥进行鉴权。
- 数据解密：针对基于区块链共享的数据进行解密操作，可以通过链上哈希码从对应的业务平台取得完整数据。
- 信息验证：需要提供对外验证接口，用于将完整信息进行链上哈希码的验真，保证共享数据的真实有效。
- 数据标准化：支持按照服务中心约定的数据交互模板进行数据的标准化格式转换，从而在异构系统之间进行数据共享。

4.6 平台化建设与运营

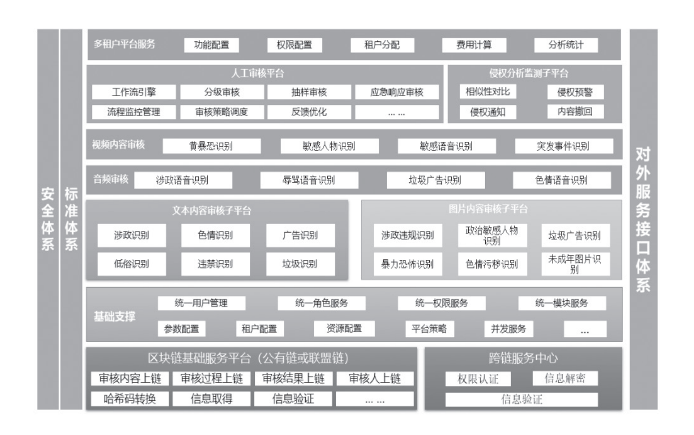
内容审核的关键核心是公信力，而区块链正是提升公信力、彰显公信力的有力技术体系，因此基于区块链建设的审核系统更易于被认同。在系统建设过程中，平台化的建设思路是一个重要的核心要点，这将为后续审核业务的扩展提供基础支撑。

通过数据分离、多租户的基础架构来保障平台化的建设，为每个用户、每种业务在区块链上按照逻辑空间分离的方式进行数据存储，功能按照模块支持灵活的组合配置，从而为平台用户量身配置各类审核能力。同时通过权限角色设置，为平台不同用户间数据隔离以及数据互通提供支持。平台化是服务的基础，服务过程是平台价值的体现也是平台迭代提升自身能力的动力源，因此利用区块链技术的优势特点，进行平台化的建设和运营是内容审核系统的建设重点。

5. 基于区块链技术的 UGC 内容风控系统功能架构

基于区块链的内容审核系统包含了区块链基础服务、跨链服务中心、平台基础功能支撑、音视频图文审核、人工审核、侵权分析检测、多租户平台服务等几个主要功能模块。

整个系统的数据核心是系统内的业务数据库以及区块链分布式存储，通过业务数据库中存储的业务数据和区块链上存储的哈希码数据相配合，来达成数据上链的同时又保证了数据量大小可控。系统通过多类型数据内容审核、人工审核以及对外服务平台来达成从审核到服务的全链条应用。同时基于标准的安全体系、标准接口服务体系来保障系统的安全可靠及对外服务的开放性。



结语

中共中央总书记习近平同志在主持中央政治局第十八次集体学习时强调，区块链技术的集成应用在新的技术革新和产业变革中起着重要作用，我们要把区块链作为核心技术自主创新的重要突破口，明确主攻方向，加大投入力度，着力攻克一批关键核心技术，加快推动区块链技术和产业创新发展。

以区块链为技术支撑的 UGC 内容审核在传统审核系统的基础上，充分发挥区块链技术特点，将审核业务流程和内容在区块链上进行存储和使用，从而提升安全、开放、共享等多方面的应用能力，并在公信力、平台化方面可以做出不断的尝试。伴随着《基于区块链的内容审核标准体系（2021 版）》标准体系的发布，内容审核领域的区块链应用将会有更多的尝试和探索，在保障内容安全生产方面也将产生更多的优秀产品和解决方案。媒

参考文献

[1] 吴迪 . 区块链技术在媒体行业应用的思考 [J]. 办公室业务，2021（14）：158−159.

[2] 杨梅 . 区块链技术推动我国融媒体发展研究 [J]. 传媒，2021（5）：38−40.

[3] 冯宇霄 . 短视频内容监管研究 [D]. 北京：北京印刷学院，2021.

[4] 纪昊仁 .AI+ 区块链在新闻报道中的应用前景 [J]. 中国传媒科技 2020（8）：32−34.

[5] 何帅，黄襄念，陈晓亮 . 区块链跨链技术发展及应用研究综述 [J]. 西华大学学报（自然科学版），2021（3）：1−14.

作者简介：张福成（1982−），男，辽宁沈阳，东软集团媒体事业中心 咨询顾问，研究方向：传媒科技；郭志帅（1976−），男，辽宁沈阳，东软集团媒体事业中心 业务总监，研究方向：传媒科技；肖勇（1978−），男，辽宁沈阳，东软集团媒体事业中心 总经理，研究方向：传媒科技。

（责任编辑：李净）